

SEQUENCE LISTING

<110> Lukyanov, Sergey
Fradkov, Arcady
Labas, Yulii
Matz, Mikhail
Lukyanov, Konstantin
Gurskaya, Nadezda

<120> FAR RED SHIFTED FLUORESCENT PROTEINS

<130> CLON-028WO

<140> Unassigned

<141> 2001-10-12

<150> 60/240,018

<151> 2000-10-12

<150> 60,306,131

<151> 2001-07-16

<160> 28

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 910

<212> DNA

<213> heteractis crispa

<400> 1

```
accatttgcct ttggttcctt ggcaaacgaa agtttagaac gaaaactgac ccaaattaca 60
tcttcctcct ggatccttac catggctggt ttgttgaaag aaagtatgcg catcaagatg 120
tacatggaag gcacgggttaa tggccattat ttcaagtgtg aaggagaggg agacggcaac 180
ccatttacag gtacgcagag catgaggatt catgtcaccg aaggggctcc attaccattt 240
gccttcgaca ttttggcacc gtgttgtgag tacggcagca ggacctttgt ccaccatacg 300
gcagagattc cegatttctt caagcagtct ttccctgaag gctttacttg ggaaagaacc 360
acaacctatg aagatggagg cattcttact gctcatcagg acacaagcct ggaggggaac 420
tgccttatat acaaggtgaa agtccttggt accaattttc ctgctgatgg ccccgatg 480
aagaacaaat caggaggatg ggagccatgc actgaggtgg tttatccaga gaatggtg 540
ctgtgtggac gtaatgtgat ggcccttaaa gtccgtgatc gtcgtttgat ctgccatctc 600
tatacttctt acaggtccaa gaaagcagtc cgtgccttga caatgccagg atttcatttt 660
acagacatcc gccttcagat gccgaggaaa acgaaagacg agtactttga actgtacgaa 720
gcactctgtg ctaggtacag tgatcttctt gaaaaagcaa attgattgtt cccagtga 780
ccagactgct gtcagctttt ggttaaagcc cgaaagacaa aaggacattt gtagtttagt 840
ttatatcttc ctttcatttg tgaatcaaca ttgtactctc tgtaaacctt taaaatgctc 900
cattaaacct
```

<210> 2

<211> 227

<212> PRT

<213> heteractis crispa

<400> 2

```
Met Ala Gly Leu Leu Lys Glu Ser Met Arg Ile Lys Met Tyr Met Glu
1           5           10          15
Gly Thr Val Asn Gly His Tyr Phe Lys Cys Glu Gly Glu Gly Asp Gly
20           25           30
```

Asn Pro Phe Thr Gly Thr Gln Ser Met Arg Ile His Val Thr Glu Gly
 35 40 45
 Ala Pro Leu Pro Phe Ala Phe Asp Ile Leu Ala Pro Cys Cys Glu Tyr
 50 55 60
 Gly Ser Arg Thr Phe Val His His Thr Ala Glu Ile Pro Asp Phe Phe
 65 70 75 80
 Lys Gln Ser Phe Pro Glu Gly Phe Thr Trp Glu Arg Thr Thr Thr Tyr
 85 90 95
 Glu Asp Gly Gly Ile Leu Thr Ala His Gln Asp Thr Ser Leu Glu Gly
 100 105 110
 Asn Cys Leu Ile Tyr Lys Val Lys Val Leu Gly Thr Asn Phe Pro Ala
 115 120 125
 Asp Gly Pro Val Met Lys Asn Lys Ser Gly Gly Trp Glu Pro Cys Thr
 130 135 140
 Glu Val Val Tyr Pro Glu Asn Gly Val Leu Cys Gly Arg Asn Val Met
 145 150 155 160
 Ala Leu Lys Val Gly Asp Arg Arg Leu Ile Cys His Leu Tyr Thr Ser
 165 170 175
 Tyr Arg Ser Lys Lys Ala Val Arg Ala Leu Thr Met Pro Gly Phe His
 180 185 190
 Phe Thr Asp Ile Arg Leu Gln Met Pro Arg Lys Thr Lys Asp Glu Tyr
 195 200 205
 Phe Glu Leu Tyr Glu Ala Ser Val Ala Arg Tyr Ser Asp Leu Pro Glu
 210 215 220
 Lys Ala Asn
 225

<210> 3
 <211> 908
 <212> DNA
 <213> heteractis crispa

<400> 3
 accatttgct ttggttcctt ggcaaacgaa agtttagacg aaaactgacc caaattacat 60
 cctcctgac cttaccatgg ctggtttggt gaaagaaagt atgcgcatca agatgtacat 120
 ggaaggcacg gttaatggcc attatttcaa gtgtgaagga gagggagacg gcaaccatt 180
 tacaggtagc cagagcatga ggattcatgt caccgaaggg gctccattac catttgcctt 240
 cgacattttg gcaccgtgtt gtgagtagcg cagcaggacc tttgtccacc ataccgcaga 300
 gattccccgat ttcttcaagc agtctttccc tgaaggcttt acttgggaaa gaaccacaac 360
 ctatgaagat ggaggcattc ttactgctca tcaggacaca agcctggagg ggaactgcct 420
 tatatacaag gtgaaagtcc ttggtaccaa tttcctgct gatggccccg tgatgaagaa 480
 caaatcagaa ggatgggagc catgcactga ggtggtttat ccagataatg gtgtcctgtg 540
 tggacgtaat gtgatggccc ttaaagtcgg tgatcgtcgt ttgatctgcc atctctatac 600
 ttcttacagg tccaagaaag cagtccgtgc cttgacaatg ccaggatttc attttacaga 660
 catccgcctt cagatgccga ggaaaacgaa agacgagtagc tttgaactgt acgaagcatc 720
 tgtggctagg tacagtgatc ttcttgaaaa agcaaattga ttgttcccag tgacaccaga 780
 ctgctgtcag cttttggtta aagccccgaaa gacaaaagga cattttagt tttagtttat 840
 attttccctt tcattttgtg aatcaacatt gtactctctg taaaccttta aaatgctcca 900
 ttaaacct 908

<210> 4
 <211> 227
 <212> PRT
 <213> heteractis crispa

<400> 4
 Met Ala Gly Leu Leu Lys Glu Ser Met Arg Ile Lys Met Tyr Met Glu
 1 5 10 15
 Gly Thr Val Asn Gly His Tyr Phe Lys Cys Glu Gly Glu Gly Asp Gly
 20 25 30

Asn Pro Phe Thr Gly Thr Gln Ser Met Arg Ile His Val Thr Glu Gly
 35 40 45
 Ala Pro Leu Pro Phe Ala Phe Asp Ile Leu Ala Pro Cys Cys Glu Tyr
 50 55 60
 Gly Ser Arg Thr Phe Val His His Thr Ala Glu Ile Pro Asp Phe Phe
 65 70 75 80
 Lys Gln Ser Phe Pro Glu Gly Phe Thr Trp Glu Arg Thr Thr Thr Tyr
 85 90 95
 Glu Asp Gly Gly Ile Leu Thr Ala His Gln Asp Thr Ser Leu Glu Gly
 100 105 110
 Asn Cys Leu Ile Tyr Lys Val Lys Val Leu Gly Thr Asn Phe Pro Ala
 115 120 125
 Asp Gly Pro Val Met Lys Asn Lys Ser Glu Gly Trp Glu Pro Cys Thr
 130 135 140
 Glu Val Val Tyr Pro Asp Asn Gly Val Leu Cys Gly Arg Asn Val Met
 145 150 155 160
 Ala Leu Lys Val Gly Asp Arg Arg Leu Ile Cys His Leu Tyr Thr Ser
 165 170 175
 Tyr Arg Ser Lys Lys Ala Val Arg Ala Leu Thr Met Pro Gly Phe His
 180 185 190
 Phe Thr Asp Ile Arg Leu Gln Met Pro Arg Lys Thr Lys Asp Glu Tyr
 195 200 205
 Phe Glu Leu Tyr Glu Ala Ser Val Ala Arg Tyr Ser Asp Leu Pro Glu
 210 215 220
 Lys Ala Asn
 225

<210> 5
 <211> 684
 <212> DNA
 <213> heteractis crispa

<400> 5
 atggctgggtt tgttgaaaga aagtatgcgc atcaagatgt acatggaagg cacggttaat 60
 ggccattatt tcaagtgtga aggagaggga gacggcaacc catttacagg tacgcagagc 120
 atgaggattc atgtcaccga aggggtcca ttaccatttg ccttcgacat tttggcaccg 180
 tgttgtgagt acggcagcag gacctttgtc caccatacgg cagagattcc cgatttcctc 240
 aagcagtctt tccctgaagg ctttacttgg gaaagaacca caacctatga agatggagggc 300
 attcttactg ctcacacagga cacaagcctg gaggggaact gccttatata caaggtgaaa 360
 gtcccttgga ccaattttcc tgctgatggc cccgtgatga agaacaatc aggaggatgg 420
 gagccaagca ctgaggtggt ttatccagag aatggtgtcc tgtgtggacg taatgtgatg 480
 gcccttaag tcggtgatcg tcgtttgatc tgccatctct ataacttctta caggtccaag 540
 aaagcagtcc gtgccttgac aatgccagga tttcatttta cagacatccg ccttcagatg 600
 ccgaggaaaa cgaaagacga gtactttgaa ctgtacgaag catctgtggc taggtacagt 660
 gatcttcctg aaaaagcaaa ttga 684

<210> 6
 <211> 227
 <212> PRT
 <213> heteractis crispa

<400> 6
 Met Ala Gly Leu Leu Lys Glu Ser Met Arg Ile Lys Met Tyr Met Glu
 1 5 10 15
 Gly Thr Val Asn Gly His Tyr Phe Lys Cys Glu Gly Glu Gly Asp Gly
 20 25 30
 Asn Pro Phe Thr Gly Thr Gln Ser Met Arg Ile His Val Thr Glu Gly
 35 40 45
 Ala Pro Leu Pro Phe Ala Phe Asp Ile Leu Ala Pro Cys Cys Glu Tyr
 50 55 60

Gly Ser Arg Thr Phe Val His His Thr Ala Glu Ile Pro Asp Phe Phe
 65 70 75 80
 Lys Gln Ser Phe Pro Glu Gly Phe Thr Trp Glu Arg Thr Thr Thr Tyr
 85 90 95
 Glu Asp Gly Gly Ile Leu Thr Ala His Gln Asp Thr Ser Leu Glu Gly
 100 105 110
 Asn Cys Leu Ile Tyr Lys Val Lys Val Leu Gly Thr Asn Phe Pro Ala
 115 120 125
 Asp Gly Pro Val Met Lys Asn Lys Ser Gly Gly Trp Glu Pro Ser Thr
 130 135 140
 Glu Val Val Tyr Pro Glu Asn Gly Val Leu Cys Gly Arg Asn Val Met
 145 150 155 160
 Ala Leu Lys Val Gly Asp Arg Arg Leu Ile Cys His Leu Tyr Thr Ser
 165 170 175
 Tyr Arg Ser Lys Lys Ala Val Arg Ala Leu Thr Met Pro Gly Phe His
 180 185 190
 Phe Thr Asp Ile Arg Leu Gln Met Pro Arg Lys Thr Lys Asp Glu Tyr
 195 200 205
 Phe Glu Leu Tyr Glu Ala Ser Val Ala Arg Tyr Ser Asp Leu Pro Glu
 210 215 220
 Lys Ala Asn
 225

<210> 7
 <211> 681
 <212> DNA
 <213> heteractis crispa

<400> 7
 tctgggtttgt tgaaagaaag tatgcgcatc aagatgtaca tggaaggcac ggttaatggc 60
 cattattttca agtgtgaagg agagggagac ggcaacccat ttgcaggtag gcagagcatg 120
 aggattcatg tcaccgaagg ggctccatta ccatttgctt tcgacatttt ggcaccgtgt 180
 tgtgagtacg gcagcaggac ctttgtccac catacggcag agattcccga tttcttcaag 240
 cagtctttcc ctgaaggctt tacttgggaa agaaccacaa cctatgaaga tggaggcatt 300
 cttactgctc atcaggacac aagcctggag gggaaactgcc ttatatacaa ggtgaaagtc 360
 cttggtacca attttcctgc tgatggcccc gtgatgaaga acaaactcagg aggatgggag 420
 ccaagcactg aggtggttta tccagagaat ggtgtcctgt gtggacgtaa tgtgatggcc 480
 cttaaagtcg gtgatcgtcg tttgatctgc catcactata cttcttacag gtccaagaaa 540
 gcagtccgtg ccttgacaat gccaggattt cattttacag acatccgcct tcagatgctg 600
 aggaaagaga aagacgagta ctttgaactg tacgaagcat ctgtgggctag gtacagtgtat 660
 cttcctgaaa aagcaaattg a 681

<210> 8
 <211> 226
 <212> PRT
 <213> heteractis crispa

<400> 8
 Ser Gly Leu Leu Lys Glu Ser Met Arg Ile Lys Met Tyr Met Glu Gly
 1 5 10 15
 Thr Val Asn Gly His Tyr Phe Lys Cys Glu Gly Glu Gly Asp Gly Asn
 20 25 30
 Pro Phe Ala Gly Thr Gln Ser Met Arg Ile His Val Thr Glu Gly Ala
 35 40 45
 Pro Leu Pro Phe Ala Phe Asp Ile Leu Ala Pro Cys Cys Glu Tyr Gly
 50 55 60
 Ser Arg Thr Phe Val His His Thr Ala Glu Ile Pro Asp Phe Phe Lys
 65 70 75 80
 Gln Ser Phe Pro Glu Gly Phe Thr Trp Glu Arg Thr Thr Thr Tyr Glu
 85 90 95

Asp Gly Gly Ile Leu Thr Ala His Gln Asp Thr Ser Leu Glu Gly Asn
 100 105 110
 Cys Leu Ile Tyr Lys Val Lys Val Leu Gly Thr Asn Phe Pro Ala Asp
 115 120 125
 Gly Pro Val Met Lys Asn Lys Ser Gly Gly Trp Glu Pro Ser Thr Glu
 130 135 140
 Val Val Tyr Pro Glu Asn Gly Val Leu Cys Gly Arg Asn Val Met Ala
 145 150 155 160
 Leu Lys Val Gly Asp Arg Arg Leu Ile Cys His His Tyr Thr Ser Tyr
 165 170 175
 Arg Ser Lys Lys Ala Val Arg Ala Leu Thr Met Pro Gly Phe His Phe
 180 185 190
 Thr Asp Ile Arg Leu Gln Met Leu Arg Lys Glu Lys Asp Glu Tyr Phe
 195 200 205
 Glu Leu Tyr Glu Ala Ser Val Ala Arg Tyr Ser Asp Leu Pro Glu Lys
 210 215 220
 Ala Asn
 225

<210> 9
 <211> 681
 <212> DNA
 <213> heteractis crispa

<400> 9
 tctggtttgt tgaagaaag tatgcgcac aagatgtaca tggaaggcac ggtaaatggc 60
 cattatttca agtgtgaagg agaggagac ggcaacccat ttgcaggtag gcagagcatg 120
 aggattcatg tcaccgaagg ggctccatta ccatttgctt tgcacatttt ggcaccgtgt 180
 tgtgcgtacg gcagcaggac ctttgtccac catacggcag agattcccga tttcttcaag 240
 cagtctttcc ctgaaggctt tacttgggaa agaaccacaa cctatgaaga tggaggcatt 300
 cttactgctc atcaggacac aagcctggag gggaactgcc ttatatacaa ggtgaaagtc 360
 cttggtacca attttcctgc tgatggcccc gtgatgaaga acaaatcagg aggatgggag 420
 ccaagcactg aggtggttta tccagagaat ggtgtcctgt gtggacgtaa tgtgatggcc 480
 cttaaagtgc gtgatcgtcg tttgatctgc catcactata cttcttacag gtccaagaaa 540
 gcagtcgctg ccttgacaat gccaggattt cattttacag acatccgcct tcagatgctg 600
 aggaaagaga aagacgagta ctttgaactg tacgaagcat ctgtggctag gtacagtgat 660
 cttcctgaaa aagcaaattg a 681

<210> 10
 <211> 226
 <212> PRT
 <213> heteractis crispa

<400> 10
 Ser Gly Leu Leu Lys Glu Ser Met Arg Ile Lys Met Tyr Met Glu Gly
 1 5 10 15
 Thr Val Asn Gly His Tyr Phe Lys Cys Glu Gly Glu Gly Asp Gly Asn
 20 25 30
 Pro Phe Ala Gly Thr Gln Ser Met Arg Ile His Val Thr Glu Gly Ala
 35 40 45
 Pro Leu Pro Phe Ala Phe Asp Ile Leu Ala Pro Cys Cys Ala Tyr Gly
 50 55 60
 Ser Arg Thr Phe Val His His Thr Ala Glu Ile Pro Asp Phe Phe Lys
 65 70 75 80
 Gln Ser Phe Pro Glu Gly Phe Thr Trp Glu Arg Thr Thr Thr Tyr Glu
 85 90 95
 Asp Gly Gly Ile Leu Thr Ala His Gln Asp Thr Ser Leu Glu Gly Asn
 100 105 110
 Cys Leu Ile Tyr Lys Val Lys Val Leu Gly Thr Asn Phe Pro Ala Asp
 115 120 125

Gly Pro Val Met Lys Asn Lys Ser Gly Gly Trp Glu Pro Ser Thr Glu
 130 135 140
 Val Val Tyr Pro Glu Asn Gly Val Leu Cys Gly Arg Asn Val Met Ala
 145 150 155 160
 Leu Lys Val Gly Asp Arg Arg Leu Ile Cys His His Tyr Thr Ser Tyr
 165 170 175
 Arg Ser Lys Lys Ala Val Arg Ala Leu Thr Met Pro Gly Phe His Phe
 180 185 190
 Thr Asp Ile Arg Leu Gln Met Leu Arg Lys Glu Lys Asp Glu Tyr Phe
 195 200 205
 Glu Leu Tyr Glu Ala Ser Val Ala Arg Tyr Ser Asp Leu Pro Glu Lys
 210 215 220
 Ala Asn
 225

<210> 11
 <211> 687
 <212> DNA
 <213> heteractis crispa

<400> 11
 atggtgagcg gcttgcgtgaa ggagagtatg cgcacatcaaga tgtacatgga gggcaccgtg 60
 aacggccact acttcaagtgc cgagggcgag ggcgacggca accccttcgc cggcaccag 120
 agcatgagaa tccacgtgac cgagggcgcc cccctgccct tcgccttcga catcctggcc 180
 ccctgctgcg agtacggcag caggaccttc gtgcaccaca ccgccgagat ccccgacttc 240
 ttcaagcaga gcttccccga gggcttcacc tgggagagaa ccaccaccta cgaggacggc 300
 ggcacatccta ccgccacca ggacaccagc ctggagggca actgcctgat ctacaagggtg 360
 aaggtgcacg gcaccaactt ccccgccgac ggcccgtga tgaagaacaa gagcggcggc 420
 tgggagccca gcaccgaggt ggtgtacccc gagaacggcg tgctgtgctg cgggaacgtg 480
 atggccctga aggtgggcga ccggcacctg atctgccacc actacaccag ctaccggagc 540
 aagaaggcgc tgcgcgccct gaccatgccc ggcttcact tcaccgacat cgggctccag 600
 atgctgcgga agaagaagga cgagtacttc gagctgtacg aggcagcgt ggcccgggtac 660
 agcgacctgc ccgagaaggc caactga 687

<210> 12
 <211> 227
 <212> PRT
 <213> heteractis crispa

<400> 12
 Met Ser Gly Leu Leu Lys Glu Ser Met Arg Ile Lys Met Tyr Met Glu
 1 5 10 15
 Gly Thr Val Asn Gly His Tyr Phe Lys Cys Glu Gly Glu Gly Asp Gly
 20 25 30
 Asn Pro Phe Ala Gly Thr Gln Ser Met Arg Ile His Val Thr Glu Gly
 35 40 45
 Ala Pro Leu Pro Phe Ala Phe Asp Ile Leu Ala Pro Cys Cys Glu Tyr
 50 55 60
 Gly Ser Arg Thr Phe Val His His Thr Ala Glu Ile Pro Asp Phe Phe
 65 70 75 80
 Lys Gln Ser Phe Pro Glu Gly Phe Thr Trp Glu Arg Thr Thr Thr Tyr
 85 90 95
 Glu Asp Gly Gly Ile Leu Thr Ala His Gln Asp Thr Ser Leu Glu Gly
 100 105 110
 Asn Cys Leu Ile Tyr Lys Val Lys Val His Gly Thr Asn Phe Pro Ala
 115 120 125
 Asp Gly Pro Val Met Lys Asn Lys Ser Gly Gly Trp Glu Pro Ser Thr
 130 135 140
 Glu Val Val Tyr Pro Glu Asn Gly Val Leu Cys Gly Arg Asn Val Met
 145 150 155 160

09976673.101204

Ala Leu Lys Val Gly Asp Arg His Leu Ile Cys His His Tyr Thr Ser
165 170 175
Tyr Arg Ser Lys Lys Ala Val Arg Ala Leu Thr Met Pro Gly Phe His
180 185 190
Phe Thr Asp Ile Arg Leu Gln Met Leu Arg Lys Thr Lys Asp Glu Tyr
195 200 205
Phe Glu Leu Tyr Glu Ala Ser Val Ala Arg Tyr Ser Asp Leu Pro Glu
210 215 220
Lys Ala Asn
225

<210> 13
<211> 687
<212> DNA
<213> heteractis crispa

<400> 13
atggtgagcg gcctgctgaa ggagagcatg cgcacatcaaga tgtacatgga gggcaccgtg 60
aacggccact acttcaagtg cgagggcgag ggcgacggca accccttcgc cggcaccag 120
agcatgcgga tccacgtgac cgagggcgcc ccctgccttgc tgcccttcga catcctggcc 180
ccctgctgag agtacggcag caggaccttc gtgcaccaca ccgccgagat ccccgacttc 240
ttcaagcaga gcttccccga gggcttcacc tgggagagaa ccaccaccta cgaggacggc 300
ggcatcctga ccgcccacca ggacaccagc ctggagggca actgcctgat ctacaagggtg 360
aagggtgctgg gcaccaactt ccccgccgac ggccccgtga tgaagaacaa gagcggcgggc 420
tgaggagccca gcaccgaggt ggtgtacccc gagaacggcg tgctgtgctg ccggaacgtg 480
atggccctga aggtggggcga ccggcggtg atctgccacc actacaccag ctaccggagc 540
aagaaggccg tgcggggccct gaccatgccg ggcttcact tcaccgacat ccggctgcag 600
atgctgcgga aggagaagga cgagtacttc gagctgtacg aggccagcgt ggcccgggtac 660
agcgacctgc ccgagaaggc caactga 687

<210> 14
<211> 227
<212> PRT
<213> heteractis crispa

<400> 14
Met Ser Gly Leu Leu Lys Glu Ser Met Arg Ile Lys Met Tyr Met Glu
1 5 10 15
Gly Thr Val Asn Gly His Tyr Phe Lys Cys Glu Gly Glu Gly Asp Gly
20 25 30
Asn Pro Phe Ala Gly Thr Gln Ser Met Arg Ile His Val Thr Glu Gly
35 40 45
Ala Pro Leu Pro Phe Ala Phe Asp Ile Leu Ala Pro Cys Cys Glu Tyr
50 55 60
Gly Ser Arg Thr Phe Val His His Thr Ala Glu Ile Pro Asp Phe Phe
65 70 75 80
Lys Gln Ser Phe Pro Glu Gly Phe Thr Trp Glu Arg Thr Thr Thr Tyr
85 90 95
Glu Asp Gly Gly Ile Leu Thr Ala His Gln Asp Thr Ser Leu Glu Gly
100 105 110
Asn Cys Leu Ile Tyr Lys Val Lys Val His Gly Thr Asn Phe Pro Ala
115 120 125
Asp Gly Pro Val Met Lys Asn Lys Ser Gly Gly Trp Glu Pro Ser Thr
130 135 140
Glu Val Val Tyr Pro Glu Asn Gly Val Leu Cys Gly Arg Asn Val Met
145 150 155 160
Ala Leu Lys Val Gly Asp Arg Arg Leu Ile Cys His His Tyr Thr Ser
165 170 175
Tyr Arg Ser Lys Lys Ala Val Arg Ala Leu Thr Met Pro Gly Phe His
180 185 190

Phe Thr Asp Ile Arg Leu Gln Met Leu Arg Lys Glu Lys Asp Glu Tyr
 195 200 205
 Phe Glu Leu Tyr Glu Ala Ser Val Ala Arg Tyr Ser Asp Leu Pro Glu
 210 215 220
 Lys Ala Asn
 225

<210> 15
 <211> 1396
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> fusion construct

<400> 15
 accggtcgcc accatggtga gcggcctgct gaaggagagc atgcgcatca agatgtacat 60
 ggagggcacc gtgaacggcc actacttcaa gtgcgagggc gagggcgacg gcaacccctt 120
 cgccggcacc cagagcatgc ggatccacgt gaccgagggc gccccctgc ccttcgcctt 180
 cgacatcctg gccccctgct gcgagtacgg cagcaggacc ttcgtgcacc acaccgccga 240
 gatccccgac ttcttcaagc agagcttccc cgagggttcc acctgggaga gaaccaccac 300
 ctacgaggac ggcggcatcc tgaccgccc aacaggacacc agcctggagg gcaactgcct 360
 gatctacaag gtgaaggtgc tgggcaccaa cttccccgcc gacggccccg tgatgaagaa 420
 caagagcggc ggctgggagc ccagcaccga ggtggtgtac cccgagaacg gcgtgctgtg 480
 cggccggaac gtgatggccc tgaaggtggg cgaccggcgg ctgatctgcc accactacac 540
 cagctaccgg agcaagaagg ccgtgcgggc cctgaccatg cccggcttcc acttcaccga 600
 catccggctg cagatgctgc ggaaggagaa ggacgagtag ttcgagctgt acgaggccag 660
 cgtggccccg tacagcgacc tgcccagaaa ggccaacaga tctcccggga tggtagcgcg 720
 cctgctgaag gagagcatgc gcatcaagat gtacatggag ggcaaccgtga acggccacta 780
 cttcaagtgc gagggcgagg gcgacggcaa ccccttcgcc ggcaaccaga gcatgcggat 840
 ccacgtgacc gagggcgccc ccctgccctt cgccttcgac atcctggccc cctgctgcga 900
 gtacggcagc aggaccttcg tgcaccacac cgccgagatc cccgacttct tcaagcagag 960
 cttccccgag ggcttcacct gggagagaac caccacctac gaggacggcg gcatcctgac 1020
 cgcccaccag gacaccagcc tggagggcaa ctgcctgatc tacaaggtga aggtgctggg 1080
 caccaacttc cccgccgacg gcccgtgat gaagaacaag agcggcggct gggagcccag 1140
 caccagagtg gtgtaccccg agaacggcgt gctgtgcggc cggaacgtga tggccctgaa 1200
 ggtgggagc cgggcgctga tctgccacca ctacaccagc taccggagca agaaggccgt 1260
 gcgggcccctg accatgcccg gcttcactt caccgacatc cggctgcaga tgctgcggaa 1320
 ggagaaggac gactacttcg agctgtacga ggccagcgtg gcccgttaca gcgacctgcc 1380
 cgagaaggcc aactga 1396

<210> 16
 <211> 460
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> fusion construct

<400> 16
 Met Val Ser Gly Leu Leu Lys Glu Ser Met Arg Ile Lys Met Tyr Met
 1 5 10 15
 Glu Gly Thr Val Asn Gly His Tyr Phe Lys Cys Glu Gly Glu Gly Asp
 20 25 30
 Gly Asn Pro Phe Ala Gly Thr Gln Ser Met Arg Ile His Val Thr Glu
 35 40 45
 Gly Ala Pro Leu Pro Phe Ala Phe Asp Ile Leu Ala Pro Cys Cys Glu
 50 55 60
 Tyr Gly Ser Arg Thr Phe Val His His Thr Ala Glu Ile Pro Asp Phe
 65 70 75 80

Phe Lys Gln Ser Phe Pro Glu Gly Phe Thr Trp Glu Arg Thr Thr Thr
 85 90 95
 Tyr Glu Asp Gly Gly Ile Leu Thr Ala His Gln Asp Thr Ser Leu Glu
 100 105 110
 Gly Asn Cys Leu Ile Tyr Lys Val Lys Val Leu Gly Thr Asn Phe Pro
 115 120 125
 Ala Asp Gly Pro Val Met Lys Asn Lys Ser Gly Gly Trp Glu Pro Ser
 130 135 140
 Thr Glu Val Val Tyr Pro Glu Asn Gly Val Leu Cys Gly Arg Asn Val
 145 150 155 160
 Met Ala Leu Lys Val Gly Asp Arg Arg Leu Ile Cys His His Tyr Thr
 165 170 175
 Ser Tyr Arg Ser Lys Lys Ala Val Arg Ala Leu Thr Met Pro Gly Phe
 180 185 190
 His Phe Thr Asp Ile Arg Leu Gln Met Leu Arg Lys Glu Lys Asp Glu
 195 200 205
 Tyr Phe Glu Leu Tyr Glu Ala Ser Val Ala Arg Tyr Ser Asp Leu Pro
 210 215 220
 Glu Lys Ala Asn Arg Ser Pro Gly Met Val Ser Gly Leu Leu Lys Glu
 225 230 235 240
 Ser Met Arg Ile Lys Met Tyr Met Glu Gly Thr Val Asn Gly His Tyr
 245 250 255
 Phe Lys Cys Glu Gly Glu Gly Asp Gly Asn Pro Phe Ala Gly Thr Gln
 260 265 270
 Ser Met Arg Ile His Val Thr Glu Gly Ala Pro Leu Pro Phe Ala Phe
 275 280 285
 Asp Ile Leu Ala Pro Cys Cys Glu Tyr Gly Ser Arg Thr Phe Val His
 290 295 300
 His Thr Ala Glu Ile Pro Asp Phe Phe Lys Gln Ser Phe Pro Glu Gly
 305 310 315 320
 Phe Thr Trp Glu Arg Thr Thr Thr Tyr Glu Asp Gly Gly Ile Leu Thr
 325 330 335
 Ala His Gln Asp Thr Ser Leu Glu Gly Asn Cys Leu Ile Tyr Lys Val
 340 345 350
 Lys Val Leu Gly Thr Asn Phe Pro Ala Asp Gly Pro Val Met Lys Asn
 355 360 365
 Lys Ser Gly Gly Trp Glu Pro Ser Thr Glu Val Val Tyr Pro Glu Asn
 370 375 380
 Gly Val Leu Cys Gly Arg Asn Val Met Ala Leu Lys Val Gly Asp Arg
 385 390 395 400
 Arg Leu Ile Cys His His Tyr Thr Ser Tyr Arg Ser Lys Lys Ala Val
 405 410 415
 Arg Ala Leu Thr Met Pro Gly Phe His Phe Thr Asp Ile Arg Leu Gln
 420 425 430
 Met Leu Arg Lys Glu Lys Asp Glu Tyr Phe Glu Leu Tyr Glu Ala Ser
 435 440 445
 Val Ala Arg Tyr Ser Asp Leu Pro Glu Lys Ala Asn
 450 455 460

<210> 17
 <211> 1424
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> fusion construct

<400> 17
 accgggtcgcc accatggtga gcggcctgct gaaggagagc atgcgcatca agatgtacat 60
 ggagggcacc gtgaacggcc actacttcaa gtgcgagggc gagggcgacg gcaaccctt 120

```

cgccggcacc cagagcatgc ggatccacgt gaccgagggc gccccctgc ccttcgcctt 180
cgacatcctg gccccctgct ggcagtagcg cagcaggacc ttcgtgcacc acaccgccga 240
gatccccgac ttcttcaagc agagcttccc cgagggttcc acctgggaga gaaccaccac 300
ctacgaggac ggcggcatcc tgaccgcca ccaggacacc agcctggagg gcaactgcct 360
gatctacaag gtgaaggtgc tgggcaccaa cttccccgcc gacggccccg tgatgaagaa 420
caagagcggc ggctgggagc ccagcaccga ggtggtgtac cccgagaacg gcgtgctgtg 480
cggccggaac gtgatggccc tgaaggtggg cgaccggcgg ctgatctgcc accactacac 540
cagctaccgg agcaagaagg ccgtgcgggc cctgaccatg cccggcttcc acttcaccga 600
catccggctg cagatgctgc ggaaggagaa ggacagtagc ttcgagctgt acgaggccag 660
cgtggcccg tacagcgacc tgcccagaaa ggccaacaga tctcccgga tggtagcg 720
cctgctgaag gagagcatgc gcatcaagat gtacatggag ggcaccgtga acggccacta 780
cttcaagtgc gagggcgagg ggcagcgcaa ccccttcgcc ggcaccaga gcatgcggat 840
ccacgtgacc gagggcgccc ccctgccctt cgcttcgac atcctggccc cctgctgcga 900
gtacggcagc aggaccttcg tgcaccacac cgccgagatc cccgacttct tcaagcagag 960
cttccccgag ggcttcacct gggagagaa caccacctac gaggacggcg gcatcctgac 1020
cgcccaccag gacaccagcc tggagggcaa ctgcctgatc tacaaggtga aggtgctggg 1080
caccaacttc cccgccgacg gccccgtgat gaagaacaag agcggcggct gggagcccag 1140
caccgaggtg gtgtaccccg agaacggcgt gctgtgcggc cggaacgtga tggccctgaa 1200
ggtgggcgac cggcggctga tctgccacca ctaccaccag tacggagca agaaggccgt 1260
gcgggccctg accatgcccg gcttccactt caccgacatc cggctgcaga tgctgcgaa 1320
ggagaaggac gactacttcg agctgtacga ggccagcgtg gcccggtaga gcgacctgcc 1380
cgagaaggcc aacagaactc gagctatgga tgatgatatc gccg 1424

```

<210> 18
 <211> 470
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> fusion construct

```

<400> 18
Met Val Ser Gly Leu Leu Lys Glu Ser Met Arg Ile Lys Met Tyr Met
  1           5           10           15
Glu Gly Thr Val Asn Gly His Tyr Phe Lys Cys Glu Gly Glu Gly Asp
          20           25           30
Gly Asn Pro Phe Ala Gly Thr Gln Ser Met Arg Ile His Val Thr Glu
          35           40           45
Gly Ala Pro Leu Pro Phe Ala Phe Asp Ile Leu Ala Pro Cys Cys Glu
          50           55           60
Tyr Gly Ser Arg Thr Phe Val His His Thr Ala Glu Ile Pro Asp Phe
          65           70           75           80
Phe Lys Gln Ser Phe Pro Glu Gly Phe Thr Trp Glu Arg Thr Thr Thr
          85           90           95
Tyr Glu Asp Gly Ile Leu Thr Ala His Gln Asp Thr Ser Leu Glu
          100          105          110
Gly Asn Cys Leu Ile Tyr Lys Val Lys Val Leu Gly Thr Asn Phe Pro
          115          120          125
Ala Asp Gly Pro Val Met Lys Asn Lys Ser Gly Gly Trp Glu Pro Ser
          130          135          140
Thr Glu Val Val Tyr Pro Glu Asn Gly Val Leu Cys Gly Arg Asn Val
          145          150          155          160
Met Ala Leu Lys Val Gly Asp Arg Arg Leu Ile Cys His His Tyr Thr
          165          170          175
Ser Tyr Arg Ser Lys Lys Ala Val Arg Ala Leu Thr Met Pro Gly Phe
          180          185          190
His Phe Thr Asp Ile Arg Leu Gln Met Leu Arg Lys Glu Lys Asp Glu
          195          200          205
Tyr Phe Glu Leu Tyr Glu Ala Ser Val Ala Arg Tyr Ser Asp Leu Pro
          210          215          220
Glu Lys Ala Asn Arg Ser Pro Gly Met Val Ser Gly Leu Leu Lys Glu

```

225		230		235		240
Ser Met Arg Ile Lys Met Tyr Met Glu Gly Thr Val Asn Gly His Tyr						
	245		250			255
Phe Lys Cys Glu Gly Glu Gly Asp Gly Asn Pro Phe Ala Gly Thr Gln						
	260		265			270
Ser Met Arg Ile His Val Thr Glu Gly Ala Pro Leu Pro Phe Ala Phe						
	275		280			285
Asp Ile Leu Ala Pro Cys Cys Glu Tyr Gly Ser Arg Thr Phe Val His						
	290		295			300
His Thr Ala Glu Ile Pro Asp Phe Phe Lys Gln Ser Phe Pro Glu Gly						
	305		310			315
Phe Thr Trp Glu Arg Thr Thr Thr Tyr Glu Asp Gly Gly Ile Leu Thr						
	325		330			335
Ala His Gln Asp Thr Ser Leu Glu Gly Asn Cys Leu Ile Tyr Lys Val						
	340		345			350
Lys Val Leu Gly Thr Asn Phe Pro Ala Asp Gly Pro Val Met Lys Asn						
	355		360			365
Lys Ser Gly Gly Trp Glu Pro Ser Thr Glu Val Val Tyr Pro Glu Asn						
	370		375			380
Gly Val Leu Cys Gly Arg Asn Val Met Ala Leu Lys Val Gly Asp Arg						
	385		390			395
Arg Leu Ile Cys His His Tyr Thr Ser Tyr Arg Ser Lys Lys Ala Val						
	405		410			415
Arg Ala Leu Thr Met Pro Gly Phe His Phe Thr Asp Ile Arg Leu Gln						
	420		425			430
Met Leu Arg Lys Glu Lys Asp Glu Tyr Phe Glu Leu Tyr Glu Ala Ser						
	435		440			445
Val Ala Arg Tyr Ser Asp Leu Pro Glu Lys Ala Asn Arg Thr Arg Ala						
	450		455			460
Met Asp Asp Asp Ile Ala						
	465		470			

<210> 19
 <211> 27
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> primer

<400> 19
 acatggatcc gctggtttgt tgaaaga

27

<210> 20
 <211> 21
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> primer

<400> 20
 acctcagtgc ttggctccca t

21

<210> 21
 <211> 21
 <212> DNA
 <213> Artificial Sequence

<220>

<223> primer

<400> 21

atgggagcca agcactgagg t

21

<210> 22

<211> 32

<212> DNA

<213> Artificial Sequence

<220>

<223> primer

<400> 22

tgacaagctt ctggtgtcac tgggaacaat ca

32

<210> 23

<211> 684

<212> DNA

<213> heteractis crispa

<400> 23

atggctgggtt tgttgaaaga aagtatgcgc atcaagatgt acatggaagg cacggttaat 60
ggccattatt tcaagtgtga aggagagga gacggcaacc catttacagg tacgcagagc 120
atgaggattc atgtcaccga aggggctcca ttaccatttg ccttcgacat tttggcaccg 180
tgttgtgagt acggcagcag gacctttgtc caccatacgg cagagattcc cgatttcttc 240
aagcagtctt tccctgaagg ctttacttgg gaaagaacca caacctatga agatggaggc 300
attcttactg ctcatcagga cacaagcctg gaggggaact gccttatata caaggtgaaa 360
gtccttggta ccaattttcc tgctgatggc cccgtgatga agaacaaatc aggaggatgg 420
gagccaagca ctgaggtggg ttatccagag aatgggtgtcc tgtgtggacg taatgtgatg 480
gcccttaaag tcggtgatcg tcgtttgatc tgccatctct atacttctta caggtccaag 540
aaagcagtcc gtgccttgac aatgccagga tttcatttta cagacatccg ccttcagatg 600
ccgaggaaaa agaaagacga gtactttgaa ctgtacgaag catctgtggc taggtacagt 660
gatcttcctg aaaaagcaaa ttga 684

<210> 24

<211> 227

<212> PRT

<213> heteractis crispa

<400> 24

Met Ala Gly Leu Leu Lys Glu Ser Met Arg Ile Lys Met Tyr Met Glu
1 5 10 15
Gly Thr Val Asn Gly His Tyr Phe Lys Cys Glu Gly Glu Gly Asp Gly
20 25 30
Asn Pro Phe Thr Gly Thr Gln Ser Met Arg Ile His Val Thr Glu Gly
35 40 45
Ala Pro Leu Pro Phe Ala Phe Asp Ile Leu Ala Pro Cys Cys Glu Tyr
50 55 60
Gly Ser Arg Thr Phe Val His His Thr Ala Glu Ile Pro Asp Phe Phe
65 70 75 80
Lys Gln Ser Phe Pro Glu Gly Phe Thr Trp Glu Arg Thr Thr Thr Tyr
85 90 95
Glu Asp Gly Gly Ile Leu Thr Ala His Gln Asp Thr Ser Leu Glu Gly
100 105 110
Asn Cys Leu Ile Tyr Lys Val Lys Val Leu Gly Thr Asn Phe Pro Ala
115 120 125
Asp Gly Pro Val Met Lys Asn Lys Ser Gly Gly Trp Glu Pro Ser Thr
130 135 140
Glu Val Val Tyr Pro Glu Asn Gly Val Leu Cys Gly Arg Asn Val Met
145 150 155 160

Ala Leu Lys Val Gly Asp Arg Arg Leu Ile Cys His Leu Tyr Thr Ser
165 170 175
Tyr Arg Ser Lys Lys Ala Val Arg Ala Leu Thr Met Pro Gly Phe His
180 185 190
Phe Thr Asp Ile Arg Leu Gln Met Pro Arg Lys Lys Lys Asp Glu Tyr
195 200 205
Phe Glu Leu Tyr Glu Ala Ser Val Ala Arg Tyr Ser Asp Leu Pro Glu
210 215 220
Lys Ala Asn
225

<210> 25
<211> 680
<212> DNA
<213> heteractis crispa

<400> 25
tctggtttgt tgaagaaag tatgcgcatc aagatgtaca tggaaggcac ggттаатggc 60
cattatttca agtgtgaagg agaggagac ggcaacccat ttgcaggtag gcagagcatg 120
aggattcatg tcaccgaagg ggctccatta ccatttgcct tcgacatttt ggcaccgtgt 180
tgtgctgacg gcagcaggac ctttgtccac catacggcag agattcccga tttcttcaag 240
cagtctttcc ctgaaggctt tacttgggaa agaaccacaa cctatgaaga tggaggcatt 300
cttactgctc atcaggacac aagcctggag gggaaactgcc ttatatacaa ggtgaaagtc 360
cttggtacca attttcctgc tgatggcccc gtgatgaaga aaaatcagga ggatgggagc 420
caagcactga ggtggtttat ccagagaatg gtgtcctgtg tggacgtaat gtgatggccc 480
ttaaagtcgg tgatcgtcgt ttgatctgcc atcactatac ttcttacagg tccaagaaag 540
cagtcctgtc cttgacaatg ccaggatttc attttacaga catccgcctt cagatgctga 600
ggaaaaagaa agacgagtac ttgaaactgt acgaagcatc tgtggctagg tacagtgatc 660
ttcctgaaaa agcaaattga 680

<210> 26
<211> 226
<212> PRT
<213> heteractis crispa

<400> 26
Ser Gly Leu Leu Lys Glu Ser Met Arg Ile Lys Met Tyr Met Glu Gly
1 5 10 15
Thr Val Asn Gly His Tyr Phe Lys Cys Glu Gly Glu Gly Asp Gly Asn
20 25 30
Pro Phe Ala Gly Thr Gln Ser Met Arg Ile His Val Thr Glu Gly Ala
35 40 45
Pro Leu Pro Phe Ala Phe Asp Ile Leu Ala Pro Cys Cys Ala Tyr Gly
50 55 60
Ser Arg Thr Phe Val His His Thr Ala Glu Ile Pro Asp Phe Phe Lys
65 70 75 80
Gln Ser Phe Pro Glu Gly Phe Thr Trp Glu Arg Thr Thr Thr Tyr Glu
85 90 95
Asp Gly Gly Ile Leu Thr Ala His Gln Asp Thr Ser Leu Glu Gly Asn
100 105 110
Cys Leu Ile Tyr Lys Val Lys Val Leu Gly Thr Asn Phe Pro Ala Asp
115 120 125
Gly Pro Val Met Lys Asn Lys Ser Gly Gly Trp Glu Pro Ser Thr Glu
130 135 140
Val Val Tyr Pro Glu Asn Gly Val Leu Cys Gly Arg Asn Val Met Ala
145 150 155 160
Leu Lys Val Gly Asp Arg Arg Leu Ile Cys His His Tyr Thr Ser Tyr
165 170 175
Arg Ser Lys Lys Ala Val Arg Ala Leu Thr Met Pro Gly Phe His Phe
180 185 190

Thr Asp Ile Arg Leu Gln Met Leu Arg Lys Lys Lys Asp Glu Tyr Phe
 195 200 205
 Glu Leu Tyr Glu Ala Ser Val Ala Arg Tyr Ser Asp Leu Pro Glu Lys
 210 215 220
 Ala Asn
 225

<210> 27
 <211> 910
 <212> DNA
 <213> heteractis crispa

<400> 27
 accatttgct ttggttcctt ggcaaacgaa agtttagaac gaaaactgac ccaaattaca 60
 tcttcctcct ggatccttac catggctggt ttgttgaaag aaagtatgag catcaagatg 120
 tacatggaag gcacgggttaa tggccattat ttcaagtgtg aaggagaggg agacggcaac 180
 ccatttacag gtacgcagag catgaggatt catgtcaccc aaggggctcc attaccattt 240
 gccttcgaca ttttggcacc gtgttgtgag tacggcagca ggacctttgt ccaccatacg 300
 gcagagattc ccgatttctt caagcagtct ttccctgaag gctttacttg ggaaagaacc 360
 acaacctatg aagatggagg cattcttact gctcatcagg acacaagcct ggaggggaac 420
 tgccttatat acaaggtgaa agtccttggg accaattttc ctgctgatgg ccccgatgatg 480
 aagaacaaat caggaggatg ggagccatgc actgaggtgg tttatccaga gaatgggtgct 540
 ctgtgtggac gtaatgtgat ggcccttaaa gtcggtgatc gtcgtttgat ctgccatctc 600
 tatacttctt acaggtccaa gaaagcagtc cgtgccttga caatgccagg atttcatttt 660
 acagacatcc gccttcagat gccgaggaaa acgaaagacg agtactttga actgtacgaa 720
 gcatctgtgg ctaggtacag tgatcttcct gaaaaagcaa attgattggt cccagtgcac 780
 ccagactgct gtcagctttt ggttaaagcc cgaaagacaa aaggacattt gtagtttagt 840
 ttatatttcc ctttcatttg tgaatcaaca ttgtactctc tgtaaaccct taaaatgctc 900
 catataacct 910

<210> 28
 <211> 227
 <212> PRT
 <213> heteractis crispa

<400> 28
 Met Ala Gly Leu Leu Lys Glu Ser Met Arg Ile Lys Met Tyr Met Glu
 1 5 10 15
 Gly Thr Val Asn Gly His Tyr Phe Lys Cys Glu Gly Glu Gly Asp Gly
 20 25 30
 Asn Pro Phe Thr Gly Thr Gln Ser Met Arg Ile His Val Thr Glu Gly
 35 40 45
 Ala Pro Leu Pro Phe Ala Phe Asp Ile Leu Ala Pro Cys Cys Glu Tyr
 50 55 60
 Gly Ser Arg Thr Phe Val His His Thr Ala Glu Ile Pro Asp Phe Phe
 65 70 75 80
 Lys Gln Ser Phe Pro Glu Gly Phe Thr Trp Glu Arg Thr Thr Thr Tyr
 85 90 95
 Glu Asp Gly Gly Ile Leu Thr Ala His Gln Asp Thr Ser Leu Glu Gly
 100 105 110
 Asn Cys Leu Ile Tyr Lys Val Lys Val Leu Gly Thr Asn Phe Pro Ala
 115 120 125
 Asp Gly Pro Val Met Lys Asn Lys Ser Gly Gly Trp Glu Pro Cys Thr
 130 135 140
 Glu Val Val Tyr Pro Glu Asn Gly Val Leu Cys Gly Arg Asn Val Met
 145 150 155 160
 Ala Leu Lys Val Gly Asp Arg Arg Leu Ile Cys His Leu Tyr Thr Ser
 165 170 175
 Tyr Arg Ser Lys Lys Ala Val Arg Ala Leu Thr Met Pro Gly Phe His
 180 185 190

Phe	Thr	Asp	Ile	Arg	Leu	Gln	Met	Pro	Arg	Lys	Thr	Lys	Asp	Glu	Tyr
		195					200					205			
Phe	Glu	Leu	Tyr	Glu	Ala	Ser	Val	Ala	Arg	Tyr	Ser	Asp	Leu	Pro	Glu
	210					215					220				
Lys	Ala	Asn													
225															